

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-85412

(43) 公開日 平成11年(1999) 3月30日

(51) Int.Cl. ⁶	識別記号	F I
G 0 6 F 3/06	3 0 5	G 0 6 F 3/06 3 0 5 C
	5 4 0	5 4 0
G 1 1 B 20/18	5 7 0	G 1 1 B 20/18 5 7 0 Z
	5 7 2	5 7 2 B
		5 7 2 F

審査請求 有 請求項の数 8 O L (全 23 頁)

(21) 出願番号 特願平9-236125

(22) 出願日 平成9年(1997) 9月1日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 浅野 良明

東京都港区芝五丁目7番1号 日本電気株式会社内

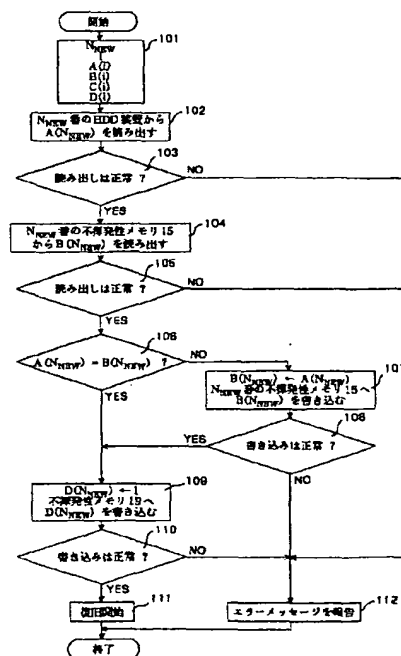
(74) 代理人 弁理士 若林 忠 (外4名)

(54) 【発明の名称】 ディスクアレイ装置制御方法およびディスクアレイ装置

(57) 【要約】

【課題】 ディスクアレイ装置の運用中に磁気ディスク装置 (HDD装置) が交換された際、HDD装置の状態を自動検出するとともに、障害HDD装置を特定する情報などを格納するための不揮発性メモリの障害も検出できるようにして、交換されたHDD装置の誤動作をなくし、HDD装置の接続ミスも検出する。

【解決手段】 障害HDD装置の交換または修理を行い、HDD装置 (NNEW番目のHDD装置) の電源を再投入した際に、そのHDD装置から、HDD装置を識別するための固有情報A (NNEW) を読み出し (ステップ102)、NNEW番目の不揮発性メモリから予め格納されている固有情報B (NNEW) を読み出し (ステップ104)、両者を比較する (ステップ106)。



【特許請求の範囲】

【請求項 1】 データを冗長構成にして格納するための複数の HDD 装置を有するディスクアレイ装置を制御するディスクアレイ装置制御方法において、前記複数の HDD 装置のそれぞれに対応する複数の第 1 の不揮発性メモリに対し、前記各 HDD 装置ごとにその HDD 装置の識別に用いる固有情報を対応する第 1 の不揮発性メモリに格納し、

前記複数の HDD 装置の状態を一括して管理するために用いる管理情報を第 2 の不揮発性メモリに格納し、前記ディスクアレイ装置の運用中に、故障した HDD 装置が交換または修理されてその HDD 装置の電源が再投入されたとき、前記交換または修理された HDD 装置に格納されている固有情報と、その HDD 装置に対応する前記第 1 の不揮発性メモリに格納されている固有情報とを比較し、それらが一致する場合には復旧を開始し、それらが一致しない場合には前記交換または修理された HDD 装置から固有情報を読み出すとともに、読み出した固有情報を前記交換または修理された HDD 装置に対応する前記第 1 の不揮発性メモリに書き込み復旧を開始し、

前記複数の HDD 装置のいずれかに障害が発生したとき、HDD 装置の復旧が開始されたとき、及び前記 HDD 装置の復旧が終了したときに、前記第 2 の不揮発性メモリ内の前記管理情報を書き換え、前記 HDD 装置から読み出した固有情報を前記第 1 の不揮発性メモリへ書き込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合には当該 HDD 装置を障害 HDD 装置として処理し、前記管理情報を前記第 2 の不揮発性メモリへ書き込む際に、書き込まれた内容と書き込んだ内容を比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する、ことを特徴とするディスクアレイ装置制御方法。

【請求項 2】 データを冗長構成にして格納するための複数の HDD 装置を有するディスクアレイ装置を制御するディスクアレイ装置制御方法において、前記 HDD 装置の識別に用いる固有情報を、前記各 HDD 装置ごとに、各 HDD 装置にそれぞれ対応する領域を有する単一の第 1 の不揮発性メモリにそれぞれ格納し、前記複数の HDD 装置の状態を一括して管理するために用いる管理情報を第 2 の不揮発性メモリに格納し、前記ディスクアレイ装置の運用中に、故障した HDD 装置が交換または修理されてその HDD 装置の電源が再投入されたとき、前記交換または修理された HDD 装置に格納されている固有情報と、前記第 1 の不揮発性メモリ内のその HDD 装置に対応する領域に格納されている固有情報とを比較し、それらが一致する場合には復旧を開始し、それらが一致しない場合には前記交換または修理された HDD 装置から固有情報を読み出すとともに、読

み出した固有情報を前記第 1 の不揮発性メモリにおける前記交換または修理された HDD 装置に対応する領域に書き込んで復旧を開始し、

前記複数の HDD 装置のいずれかに障害が発生したとき、HDD 装置の復旧が開始されたとき、及び前記 HDD 装置の復旧が終了したときに、前記第 2 の不揮発性メモリ内の前記管理情報を書き換え、

前記 HDD 装置から読み出した固有情報を前記第 1 の不揮発性メモリへ書き込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合にはディスクアレイ装置障害として処理し、

前記管理情報を前記第 2 の不揮発性メモリへ書き込む際に、書き込まれた内容と書き込んだ内容を比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する、ことを特徴とするディスクアレイ装置制御方法。

【請求項 3】 データを冗長構成にして格納するための複数の HDD 装置を有し、それぞれの冗長を構成する前記 HDD 装置のまとまりとしての論理ドライブが設定されたディスクアレイ装置を制御するディスクアレイ装置制御方法において、

前記複数の HDD 装置のそれぞれに対応する複数の第 1 の不揮発性メモリに対し、前記各 HDD 装置ごとにその HDD 装置の識別に用いる固有情報を対応する第 1 の不揮発性メモリに格納し、

前記複数の HDD 装置の状態を一括して管理するために用いる管理情報を第 2 の不揮発性メモリに格納し、

前記ディスクアレイ装置の運用中に、故障した HDD 装置が交換または修理されてその HDD 装置の電源が再投入されたとき、前記交換または修理された HDD 装置から読み出される固有情報と全ての前記第 1 の不揮発性メモリに格納されている固有情報とを比較し、一致するものがなければ復旧を開始し、一致するものがあれば前記各論理ドライブに 2 カ所以上の障害 HDD 装置がないことを確認し、前記各論理ドライブに障害 HDD 装置が 2 カ所以上ない場合には復旧を開始し、前記論理ドライブのいずれかに 2 カ所以上の障害 HDD 装置がある場合には、前記交換または修理された HDD 装置の固有情報が一致した前記第 1 の不揮発性メモリが存在する論理ドライブに 2 カ所以上の障害 HDD 装置が存在しなければ復旧を開始し、2 カ所以上の障害 HDD 装置が存在する場合には、復旧は行わないようにし、

前記複数の HDD 装置のいずれかに障害が発生したとき、HDD 装置の復旧が開始されたとき、及び前記 HDD 装置の復旧が終了したときに、前記第 2 の不揮発性メモリ内の前記管理情報を書き換え、

前記 HDD 装置から読み出した固有情報を前記第 1 の不揮発性メモリへ書き込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合には当該 HDD 装置を障害 HDD 装置として処理し、

前記管理情報を前記第2の不揮発性メモリに書込む際に、書き込まれた内容と書き込んだ内容を比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する、ことを特徴とするディスクアレイ装置制御方法。

【請求項4】 データを冗長構成にして格納するための複数のHDD装置を有し、それぞれの冗長を構成する前記HDD装置のまとまりとしての論理ドライブが設定されたディスクアレイ装置を制御するディスクアレイ装置制御方法において、

前記HDD装置の識別に用いる固有情報を、前記各HDD装置ごとに、前記各HDD装置にそれぞれ対応する記憶領域を有する単一の第1の不揮発性メモリにそれぞれ格納し、

前記複数のHDD装置の状態を一括して管理するために用いる管理情報を第2の不揮発性メモリに格納し、

前記ディスクアレイ装置の運用中に、故障したHDD装置が交換または修理されてそのHDD装置の電源が再投入されたとき、前記交換または修理されたHDD装置に格納された固有情報と前記第1の不揮発性メモリに格納されている全ての固有情報とを比較し、一致するものがないければ復旧を開始し、一致するものがあれば前記各論理ドライブに2カ所以上の障害HDD装置がないことを確認し、前記各論理ドライブに障害HDD装置が2カ所以上ない場合には復旧を開始し、前記論理ドライブのいずれかに2カ所以上の障害HDD装置がある場合には、前記交換または修理されたHDD装置の固有情報が一致した前記第1の不揮発性メモリが存在する論理ドライブに2カ所以上の障害HDD装置が存在しなければ復旧を開始し、2カ所以上の障害HDD装置が存在する場合には、復旧は行わないようにし、

前記複数のHDD装置のいずれかに障害が発生したとき、HDD装置の復旧が開始されたとき、及び前記HDD装置の復旧が終了したときに、前記第2の不揮発性メモリ内の前記管理情報を書き換え、

前記HDD装置から読み出した固有情報を前記第1の不揮発性メモリに書込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合にはディスクアレイ装置障害として処理し、

前記管理情報を前記第2の不揮発性メモリへ書込む際に、書き込まれた内容と書き込んだ内容を比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する、ことを特徴とするディスクアレイ装置制御方法。

【請求項5】 データを冗長構成にして格納するための複数のHDD装置と、

前記各HDD装置ごとに設けられ、それぞれが対応するHDD装置の識別に用いる固有情報を格納する、複数の第1の不揮発性メモリと、

前記複数のHDD装置の状態を一括して管理するために

用いる管理情報を格納するための第2の不揮発性メモリと、

前記ディスクアレイ装置の運用中に、故障したHDD装置が交換または修理されてそのHDD装置の電源が再投入されたとき、前記交換または修理されたHDD装置から読み出された固有情報と、前記交換または修理されたHDD装置に対応する前記第1の不揮発性メモリに格納されている固有情報とを比較し、それらが一致する場合には復旧を開始し、それらが一致しない場合には前記交換または修理されたHDD装置から固有情報を読み出すとともに、読み出した固有情報を前記交換または修理されたHDD装置に対応する前記第1の不揮発性メモリに書き込み、復旧を開始するHDD装置自動復旧開始手段と、

前記HDD装置から固有情報を読み出し、当該HDD装置に対応する前記第1の不揮発性メモリに書き込む固有情報書込手段と、

前記複数のHDD装置のいずれかに障害が発生したとき、前記HDD装置の復旧が開始されたとき、前記HDD装置の復旧が終了したときに、前記管理情報の内容を書き換える管理情報書換手段と、

前記HDD装置から読み出した固有情報を前記第1の不揮発性メモリへ書込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合には当該HDD装置を障害HDD装置として処理する固有情報書込障害検出手段と、

前記管理情報を前記第2の不揮発性メモリに書込む際に、書き込まれた内容と書き込んだ内容を比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する管理情報書込障害検出手段と、を具備するディスクアレイ装置。

【請求項6】 データを冗長構成にして格納するための複数のHDD装置と、

HDD装置の識別に用いる固有情報を格納するため、前記複数のHDD装置に対して共通に設けられ、それぞれのHDD装置ごとに対応する領域を有する単一の第1の不揮発性メモリと、

複数のHDD装置の状態を一括して管理するための用いる管理情報を格納するための第2の不揮発性メモリと、

前記ディスクアレイ装置の運用中に、故障したHDD装置が交換または修理されてそのHDD装置の電源が再投入されたとき、前記交換または修理されたHDD装置に格納されている固有情報と、前記第1の不揮発性メモリ内のそのHDD装置に対応する領域に格納されている固有情報とを比較し、それらが一致する場合には復旧を開始し、それらが一致しない場合には前記交換または修理されたHDD装置から固有情報を読み出すとともに、読み出した固有情報を前記第1の不揮発性メモリにおける前記交換または修理されたHDD装置に対応する領域に書き込んで復旧を開始するHDD装置自動復旧開始手段

と、
前記HDD装置から固有情報を読み出し、前記第1の不揮発性メモリ内の当該HDD装置に対応する領域に書き込む固有情報書込手段と、

前記複数のHDD装置のいずれかに障害が発生したとき、HDD装置の復旧が開始されたとき、及び前記HDD装置の復旧が終了したときに、前記管理情報を書き換える管理情報書換手段と、

前記HDD装置から読み出した固有情報を前記第1の不揮発性メモリへ書込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する不揮発性メモリ情報書込障害検出手段と、

前記管理情報を前記第2の不揮発性メモリへ書込む際に、書き込まれた内容と書き込んだ内容を比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する管理情報書込障害検出手段と、を具備するディスクアレイ装置。

【請求項7】 データを冗長構成にして格納するための複数のHDD装置と、

前記各HDD装置ごとに設けられ、それぞれが対応するHDD装置の識別に用いる固有情報を格納する、複数の第1の不揮発性メモリと、

前記複数のHDD装置の状態を一括して管理するために用いる管理情報を格納するための第2の不揮発性メモリと、

前記ディスクアレイ装置の運用中に、故障したHDD装置が交換または修理されてそのHDD装置の電源が再投入されたとき、前記交換または修理されたHDD装置から読み出される固有情報と全ての前記第1の不揮発性メモリに格納されている固有情報とを比較し、一致するものがなければ復旧を開始し、一致するものがあれば各論理ドライブに2カ所以上の障害HDD装置がないことを確認し、前記各論理ドライブに障害HDD装置が2カ所以上ない場合には復旧を開始し、前記論理ドライブのいずれかに2カ所以上の障害HDD装置がある場合には、前記交換または修理されたHDD装置の固有情報が一致した前記第1の不揮発性メモリが存在する論理ドライブに2カ所以上の障害HDD装置が存在しなければ復旧を開始し、2カ所以上の障害HDD装置が存在する場合には、復旧は行わないようにする交換HDD装置復旧開始判定手段と、

前記HDD装置から固有情報を読み出し、当該HDD装置に対応する前記第1の不揮発性メモリに書き込む固有情報書込手段と、

前記複数のHDD装置のいずれかに障害が発生したとき、前記HDD装置の復旧が開始されたとき、前記HDD装置の復旧が終了したときに、前記管理情報の内容を書き換える管理情報書換手段と、

前記HDD装置から読み出した固有情報を前記第1の不

揮発性メモリへ書込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合には当該HDD装置を障害HDD装置として処理する固有情報書込障害検出手段と、

前記管理情報を前記第2の不揮発性メモリに書込む際に、書き込まれた内容と書き込んだ内容を比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する管理情報書込障害検出手段と、を具備し、それぞれの冗長を構成する前記HDD装置のまとまりとして、前記論理ドライブが設定されているディスクアレイ装置。

【請求項8】 データを冗長構成にして格納するための複数のHDD装置と、

HDD装置の識別に用いる固有情報を格納するため、前記複数のHDD装置に対して共通に設けられ、それぞれのHDD装置ごとに対応する領域を有する単一の第1の不揮発性メモリと、

複数のHDD装置の状態を一括して管理するための用いる管理情報を格納するための第2の不揮発性メモリと、

前記ディスクアレイ装置の運用中に、故障したHDD装置が交換または修理されてそのHDD装置の電源が再投入されたとき、前記交換または修理されたHDD装置に格納された固有情報と前記第1の不揮発性メモリに格納されている全ての固有情報とを比較し、一致するものがなければ復旧を開始し、一致するものがあれば各論理ドライブに2カ所以上の障害HDD装置がないことを確認し、前記各論理ドライブに障害HDD装置が2カ所以上ない場合には復旧を開始し、前記論理ドライブのいずれかに2カ所以上の障害HDD装置がある場合には、前記交換または修理されたHDD装置の固有情報が一致した前記第1の不揮発性メモリが存在する論理ドライブに2カ所以上の障害HDD装置が存在しなければ復旧を開始し、2カ所以上の障害HDD装置が存在する場合には、復旧は行わないようにする交換HDD装置復旧開始判定手段と、

前記HDD装置から固有情報を読み出し、前記第1の不揮発性メモリ内の当該HDD装置に対応する領域に書き込む固有情報書込手段と、

前記複数のHDD装置のいずれかに障害が発生したとき、HDD装置の復旧が開始されたとき、及び前記HDD装置の復旧が終了したときに、前記管理情報を書き換える管理情報書換手段と、

前記HDD装置から読み出した固有情報を前記第1の不揮発性メモリへ書込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する不揮発性メモリ情報書込障害検出手段と、

前記管理情報を前記第2の不揮発性メモリへ書込む際に、書き込まれた内容と書き込んだ内容を比較し、それらが一致していない場合にはディスクアレイ装置障害と

して処理する管理情報書込障害検出手段と、を具備し、それぞれの冗長を構成する前記HDD装置のまとまりとして、前記論理ドライブが設定されているディスクアレイ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ディスクアレイ装置に関し、特に、データを冗長構成にして複数の磁気ディスク装置（HDD装置）に格納するディスクアレイ装置制御方法及びディスクアレイ装置に関する。

【0002】

【従来の技術】サーバ機などでは、信頼性や耐故障性の向上、データアクセス速度の向上を目的として、複数の磁気ディスク装置（HDD装置）からなり、データを分解してこれら複数の磁気ディスク装置に対して並列にリード/ライトを行えるようにしたディスクアレイ装置が用いられる。ディスクアレイ装置では、障害を起こしたHDD装置を特定する情報を保持できるようになっており、障害発生時などにはこの情報を利用して、障害からの復旧が行えるようになっている。

【0003】従来のディスクアレイ装置では、障害HDD装置を特定する情報を揮発性メモリに格納していたため、停電等によるシステムダウンが発生したとき、障害HDD装置を特定する情報が失われていた。このため、障害HDD装置を特定する情報を不揮発性メモリに格納するディスクアレイ装置が各種提案されている。例えば、本願発明者は、既に特願平9-56234号において、磁気ディスクの状態記憶に不揮発性メモリを用いるとともに、各HDD装置の状態を立上げ時に自動検出し、また、不揮発性メモリ自体の障害やHDD装置の接続ミスも検出できるディスクアレイシステムを提案した。

【0004】図5に、この従来のディスクアレイシステムの構成を示す。ディスクアレイシステムは、上位装置から送信されたデータを冗長構成のものとして格納する複数（ここでは例えば5台とする）の磁気ディスク装置（HDD装置）541～545と、この複数のHDD装置541～545を識別する固有情報を各HDD装置ごとに記憶する不揮発性のメモリ（不揮発性メモリ）551～555と、この不揮発性メモリ551～555に格納された固有情報に基づいてHDD装置541～545の接続状態を判定するとともにHDD装置541～545に異常が発生したときに冗長構成に基づいて当該異常により失われたデータを復旧させる磁気ディスク制御装置（ディスクアレイコントローラ）3と、ディスクアレイコントローラ3内の後述する状態情報制御手段6によって編集された状態情報を記憶する不揮発性の状態情報記憶部19と、を有している。

【0005】ディスクアレイコントローラ3には、ディスクアレイコントローラ3がHDD装置541～545の初期化する時に当該各HDD装置541～545から固有

情報を読み出すとともに当該各固有情報をそれぞれの不揮発性メモリ551～555へ格納する初期化制御手段5と、HDD装置541～545の障害発生の有無および復旧処理中か否かの動作状態を管理する状態情報をHDD装置541～545の状態の変化に応じて書き換える状態情報制御手段6と、ディスクアレイコントローラ3がHDD装置541～545を再立ち上げをするときに状態情報記憶部19に格納された状態情報に基づいて、当該再立ち上げ直前の各HDD装置541～545の動作状態を再現する直前状態再現手段7と、不揮発性メモリ551～555およびHDD装置541～545にそれぞれ格納された固有情報を比較するとともに当該比較結果に基づいて再立ち上げから当該HDD装置541～545が入れ替えられたか否かの接続状態を判定する接続状態判定手段8と、この接続状態判定手段8によって判定された接続状態情報と直前状態再現手段7によって再現された直前の各HDD装置541～545の動作状態情報とに基づいて復旧処理の継続または開始を制御する復旧制御手段9とが設けられている。

【0006】この図5に示すディスクアレイシステムでは、初期化制御手段5が、HDD装置に固有な情報を不揮発性メモリ551～555に格納するため、接続状態判定手段8は、不揮発性メモリ551～555およびHDD装置541～545にそれぞれ格納された固有情報を比較するとともに当該比較結果に基づいて再立ち上げ直前から当該HDD装置541～545が入れ替えられたか否かの接続状態を判定することができる。すなわち、HDD装置が交換された場合、不揮発性メモリに格納された固有情報とその交換されたHDD装置の固有情報は必ず異なり、一方、交換されていない場合には必ず一致するため、接続状態判定手段8は、現在のHDD装置の接続状態を正確に判定することができる。さらに、このディスクアレイシステムでは、状態情報制御手段6が、HDD装置541～545の障害発生の有無および復旧処理中か否かの動作状態を管理する状態管理情報をHDD装置541～545の変化に応じて書き換え、この状態情報を不揮発性の状態情報記憶部19に格納する。このため、停電等によるシステムダウンが生じても、HDD装置541～545を再立ち上げするときに、直前状態再現手段7は、この状態情報記憶部19に格納された状態情報に基づいて当該再立ち上げ直前の各HDD装置541～545の動作状態を再現することができる。すなわち、システムダウン等の不測の事故が生じても、「正常」、「障害」、「復旧処理中」等の直前の各HDD装置の動作状態が再現される。

【0007】次に、図5に示す従来のディスクアレイシステムにおける再起動時の動作の概略を図6乃至図10のフローチャートを参照して説明する。図6乃至図10は、全体で1枚のフローチャートを構成するものであり、丸付きの符号A～Hにより、相互の接続関係が示さ

れている。

【0008】まず、カウンタ*i*に“1”、障害HDD装置台数NEERに“0”、障害HDD識別パラメータNdiskに“0”、不揮発性メモリ19に格納されている障害状態を示す管理情報配列C(*i*)の各要素に“0”、各不揮発性メモリ55から読み出した固有情報の配列E(*i*)の各要素に“0”、不揮発性メモリ19に格納されている復旧状態を示す管理情報配列F(*i*)の各要素に“0”をセットする(ステップ401)。そして、不揮発性メモリ19から、各HDD装置の状態を示す管理情報配列C(*i*)と、復旧状態を示す管理情報配列F(*i*)とを読み出し(ステップ402)、読み出しが正常に行えたかどうかを判別する(ステップ403)。

【0009】ステップ403において読み出しが正常に行えなかった場合には、ディスクアレイシステムの障害としてエラーメッセージを出力し(ステップ516)、処理を終了する。一方、ステップ403において読み出しが正常に行えた場合には、C(*i*)が“1”かどうかを判別し(ステップ404)、C(*i*)が“1”であれば、図7に示すステップ415に処理を移行し、*i*番のHDD装置から固有情報を読み出してD(*i*)に格納し、さらに、読み出しが正常に行えたかどうかを判別する(ステップ417)。ここで読み出しが正常でなければステップ420に進み、一方、ステップ417で読み出しが正常であれば、*i*番の不揮発性メモリからE(*i*)を読み出し(ステップ416)、ステップ416での読み出しが正常かどうかを判別する(ステップ418)。ステップ418において、読み出しが正常でなければステップ420へ進み、読み出しが正常であればステップ412へ進む。

【0010】ステップ420では、*i*番のHDD装置から読み出した固有情報D(*i*)と*i*番の不揮発性メモリから読み出した固有情報が異なるように、固有情報の一部の書き換えを行い、その後、ステップ412へ進む。例えば、D(*i*)に“FFh”(hは16進表示であることを示す)を代入し、E(*i*)に“00h”を代入する。このステップ420での固有情報の書き換えにより、*i*番の不揮発性メモリ55に異常があることを記憶する。また、ステップ412では、NEERに1を加算、Ndiskに*i*をセットし、図8のステップ413へ進む。

【0011】図6に示すステップ404においてC(*i*)が1でなければ、すなわち、再立ち上げ直前に障害なしであったHDD装置については、図8に示すステップ405に処理を移行し、この*i*番のHDD装置から固有情報D(*i*)を読み出し、この読み出しが正常に行えたかどうかを判別する(ステップ406)。ここで読み出しが正常に行えなかった場合にはステップ419に進み、読み出しが正常に行えた場合には*i*番の不揮発性メモリ55から格納されている固有情報E(*i*)を読み出し(ステップ407)、固有情報E(*i*)の読み出しが正常に行え

たかどうかを判別する(ステップ408)。ステップ408において、読み出しが正常に行えなかった場合には、ステップ419に進む。ステップ419では、HDD装置の異常またはその不揮発性メモリ55に異常がある場合には、*i*番のHDD装置から読み出した固有情報D(*i*)と*i*番の不揮発性メモリから読み出した固有情報が異なるように、固有情報の一部の書き換えを行い、その後、ステップ410へ進む。例えば、D(*i*)に“FFh”を代入し、E(*i*)に“00h”を代入する。

【0012】ステップ408において固有情報E(*i*)の読み出しが正常に行えた場合には、*i*番のHDD装置から読み出した固有情報D(*i*)と*i*番の不揮発性メモリ55から読み出した固有情報E(*i*)を比較し、D(*i*)とE(*i*)が同じか異なるかを判別する(ステップ409)。ステップ409においてD(*i*)≠E(*i*)であればステップ410へ進み、D(*i*)=E(*i*)であればステップ413へ進む。ステップ409においてD(*i*)とE(*i*)が一致しないのは、対応するHDD装置が交換されたことを意味する。このため、ステップ410では、NEERに“1”を加算、Ndiskに*i*をセット、C(*i*)に“1”をセットし、管理情報を格納するための不揮発性メモリ19にC(*i*)を書き込む。そして、この書き込みが正常に行えたかどうかを判別する(ステップ411)。書き込みが正常でない場合には、ステップ412においてエラーメッセージを出力してから処理を終了し、書き込みが正常である場合には、ステップ413へ進む。

【0013】ステップ413では、次のHDD装置について上述と同様の繰り返すため、パラメータ*i*に“1”を加算し、加算後のパラメータ*i*が全HDD装置の台数以下であるかを判別する(ステップ414)。*i*が全HDD装置の台数以下の場合には、図6に示すステップ404へ戻り、次のHDD装置の固有情報と不揮発性メモリ55の固有情報との比較が行われる。このような動作を*i*が全HDD装置台数より大きくなるまで繰り返すことにより、NEERには、正常でないHDD装置の数がセットされ、管理情報C(*i*)には、障害の発生しているHDD装置を識別するビットがセットされることになる。また、管理情報を格納する不揮発性メモリ55の障害によるディスクアレイシステム障害が検出されたことになる。

【0014】なお、上述した例では、ステップ419およびステップ420において、D(*i*)に“FFh”を、E(*i*)に“00h”をそれぞれ書き込んでいるが、要はD(*i*)とE(*i*)とが異なるようにすればよいので、書き込む値はこれらに限られるものではない。これは、読み出しには失敗するがデータは正常なものが送られてくる場合があるためで、回路の設計方法やデータの格納方法によっては使用しなくても構わない場合もある。以上の図6乃至図8に示す処理により、現在のHDD装置の接続状態が確認され、HDD装置が交換されている場合に

はその交換されたHDD装置の番号がNdiskに格納されることになる。この接続状態の確認は、接続状態判定手段8により行われる。

【0015】全HDD装置について上述の処理が行われた場合、すなわち、ステップ414でiが全HDD装置台数を上回る場合には、ステップ501に移行し、図9及び図10に示す処理を行う。

【0016】まず、ステップ501において、障害HDD装置台数カウンタNEERの判定を行い、NEERが“0”であるときには、全HDD装置が正常であるので、立ち上げ動作を終了する。NEERが“2”以上であるときには、ここでのディスクレイシステムの冗長構成ではカバーできない障害であるため、障害ディスクアレイ装置とし、エラーメッセージを出力して（ステップ516）、処理を終了する。NEERが“1”であるときには、復旧状態を示す管理情報F(Ndisk)から、電源切断前の状態が復旧中であるかを判断する（ステップ502）。ステップ502において、F(Ndisk)=1すなわち電源切断前の状態が復旧中であれば、次に、HDD装置から読み出した固有情報D(Ndisk)と不揮発性メモリ55から読み出した固有情報E(Ndisk)が同じであるかを比較する（ステップ503）。この比較により、D(Ndisk)=E(Ndisk)すなわちF(Ndisk)で特定されるHDD装置が交換されていないと判断された場合には、ステップ518以下の自動復旧処理を行う。一方、D(Ndisk)とE(Ndisk)が異なる場合には、当該HDD装置を使用不可として（ステップ517）、処理を終了する。

【0017】一方、ステップ502においてF(Ndisk)≠1すなわち電源切断前が復旧中でない場合には、HDD装置から読み出した固有情報D(Ndisk)と不揮発性メモリ55から読み出した固有情報E(Ndisk)が同じであるかを比較する（ステップ504）。ステップ504においてD(Ndisk)=E(Ndisk)であるときは、一旦障害有りと判定されたHDD装置であるため、再度、このNdisk番のHDD装置が正常か否かのテストを行う（ステップ506）。このテスト結果が正常であるかどうかを判別し（ステップ507）、正常でなければ、ステップ517に移行し、当該HDD装置を使用不可能として処理を終了する。一方、ステップ507においてNdisk番のHDD装置のテスト結果が正常であれば、復旧状態を示す管理情報配列F(Ndisk)に“1”を格納し、不揮発性メモリ19にこのF(Ndisk)を書き込み（ステップ514）、不揮発性メモリ19への書き込みが正常に行えたかどうかを判別する（ステップ515）。ステップ515において書き込みが正常に行えていない場合には、エラーメッセージを出力して（ステップ516）、処理を終了し、書き込みが正常に行われている場合には、ステップ518に移行して、自動復旧を行う。

【0018】ステップ518では、Ndisk番のHDD装

置に他のHDD装置内のデータから再構築されたデータを書き込むことによって、このNdisk番のHDD装置の復旧作業を行う。この復旧作業が終了したら、管理情報C(Ndisk)、D(Ndisk)にそれぞれ“0”をセットし（ステップ519）、これによって再立ち上げ動作を終了する。

【0019】また、ステップ504において、HDD装置から読み出した固有情報D(Ndisk)と不揮発性メモリ55から読み出した固有情報E(Ndisk)が異なる場合、すなわち、障害処理中ではないHDD装置が交換された場合には、図10のステップ505に処理を移行し、上位装置からの復旧命令を待つ。そして、復旧命令を受信したら、Ndisk番のHDD装置のテストを行う（ステップ508）。ステップ508でのHDD装置のテストの結果が正常かどうかを判別し（ステップ509）、正常でない場合には、当該HDD装置を使用不可とするためにステップ517に移行し、HDD装置のテストが正常な場合には、Ndisk番のHDD装置からD(Ndisk)を読み出す（ステップ510）。ステップ510での読み出しが正常に行えたかどうかを判別し（ステップ511）、読み出しが正常に行えなかった場合には、ステップ517へ進み、読み出しが正常に行えた場合には、Ndisk番の不揮発性メモリ55へD(Ndisk)を書き込み（ステップ512）、不揮発性メモリ55への書き込みが正常に行えたかどうかを判別する（ステップ513）。ステップ513において、不揮発性メモリ55へD(Ndisk)が正常に書き込めなかったと判別された場合には、ステップ517へ進み、正常に書き込んだ場合には、復旧状態を示す管理情報F(Ndisk)に“1”をセットして不揮発性メモリ19へF(Ndisk)を書き込むために、上述のステップ514に進む。

【0020】以上のようにして、図5に示す従来のディスクアレイシステムの再起動処理が行われる。

【0021】

【発明が解決しようとする課題】しかしながら、上述した従来のディスクアレイシステムでは、障害HDD装置を取り替えたときは、一旦ディスクアレイ装置の電源を落とし、そののちディスクアレイ装置全体の再立ち上げを行うことにより、復旧を開始する。しかしながら、ディスクアレイ装置の多くは論理ドライブを複数個有し、また、24時間連続運用されている場合があるため、このような場合には、上述のディスクアレイシステムでは、1つの論理ドライブに障害HDD装置が発生してもディスクアレイ装置の電源を落すことができないことがあり、結果として、障害HDD装置の復旧を行えない、という不都合が生じる。

【0022】また、1つのディスクアレイ装置内で障害HDD装置が複数の論理ドライブ内に存在し、それらの障害HDD装置の交換を行う際、障害HDD装置の他に誤って正常なHDD装置を抜き取る場合が考えられる。

誤って抜き取ったHDD装置は、修理や検査において正常であることからディスクアレイ装置の元の位置に戻そうとするが、このとき誤って抜き取った正常なHDD装置が存在する論理ドライブに他の障害HDD装置が存在しない場合には問題ないが、他の障害HDD装置が存在する場合には、冗長構成がなくなることになる。このような場合、誤って抜き取ったHDD装置のデータが変わっていなければ、人手介入により、データの復旧を行うことが可能であるが、このHDD装置を他の論理ドライブHDD装置の搭載位置に乗せた場合、HDD装置は正常に動作するので復旧を開始したら、もともと正常なHDD装置が存在した論理ドライブの冗長構成がなくなり、人手介入による復旧も不可能になる、という不都合が生じる。

【0023】本発明は、従来のディスクアレイ装置が有するかかる不都合を改善することを目的とし、特に、ディスクアレイ装置の運用中に、状態に応じて適切に復旧処理を行うことのできるディスクアレイ装置を提供することを目的とする。

【0024】具体的には、ディスクアレイ装置運用中にHDD装置の交換が行われた際、HDD装置の状態を自動検出するとともに不揮発性メモリの障害も検出することにより、交換されたHDD装置の誤動作が生じることがなくなるようにして、システムの信頼性の向上を図る。また、HDD装置の接続ミスも検出することにより、人為的ミスによりデータが失われることがなくなるようにしてシステムの信頼性の向上を図る。

【0025】

【課題を解決するための手段】本発明の第1のディスクアレイ装置制御方法は、データを冗長構成にして格納するための複数のHDD装置を有するディスクアレイ装置を制御するディスクアレイ装置制御方法において、複数のHDD装置のそれぞれに対応する複数の第1の不揮発性メモリに対し、各HDD装置ごとにそのHDD装置の識別に用いる固有情報を対応する第1の不揮発性メモリに格納し、複数のHDD装置の状態を一括して管理するために用いる管理情報を第2の不揮発性メモリに格納し、ディスクアレイ装置の運用中に、故障したHDD装置が交換または修理されてそのHDD装置の電源が再投入されたとき、交換または修理されたHDD装置に格納されている固有情報と、そのHDD装置に対応する第1の不揮発性メモリに格納されている固有情報とを比較し、それらが一致する場合には復旧を開始し、それらが一致しない場合には交換または修理されたHDD装置から固有情報を読み出すとともに、読み出した固有情報を交換または修理されたHDD装置に対応する第1の不揮発性メモリに書き込み復旧を開始し、複数のHDD装置のいずれかに障害が発生したとき、HDD装置の復旧が開始されたとき、及びHDD装置の復旧が終了したときに、第2の不揮発性メモリ内の管理情報を書き換え、H

DD装置から読み出した固有情報を第1の不揮発性メモリへ書き込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合には当該HDD装置を障害HDD装置として処理し、管理情報を第2の不揮発性メモリへ書き込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する。

【0026】本発明の第2のディスクアレイ装置制御方法は、上述の第1のディスクアレイ装置制御装置においてHDD装置ごとに独立に設けられた第1の不揮発性メモリに固有情報を格納する代わりに、各HDD装置ごとにその固有情報を、各HDD装置にそれぞれ対応する領域を有する単一の第1の不揮発性メモリにそれぞれ格納するようにし、また、HDD装置から読み出した固有情報を第1の不揮発性メモリへ書き込む際に書き込まれた内容と書き込んだ内容が一致していない場合にはディスクアレイ装置障害として処理するようにしたものである。

【0027】本発明の第3のディスクアレイ装置制御方法は、データを冗長構成にして格納するための複数のHDD装置を有し、それぞれの冗長を構成するHDD装置のまとまりとしての論理ドライブが設定されたディスクアレイ装置を制御するディスクアレイ装置制御方法において、複数のHDD装置のそれぞれに対応する複数の第1の不揮発性メモリに対し、各HDD装置ごとにそのHDD装置の識別に用いる固有情報を対応する第1の不揮発性メモリに格納し、複数のHDD装置の状態を一括して管理するために用いる管理情報を第2の不揮発性メモリに格納し、ディスクアレイ装置の運用中に、故障したHDD装置が交換または修理されてそのHDD装置の電源が再投入されたとき、交換または修理されたHDD装置から読み出される固有情報と全ての第1の不揮発性メモリに格納されている固有情報とを比較し、一致するものがなければ復旧を開始し、一致するものがあれば各論理ドライブに2カ所以上の障害HDD装置がないことを確認し、各論理ドライブに障害HDD装置が2カ所以上ない場合には復旧を開始し、論理ドライブのいずれかに2カ所以上の障害HDD装置がある場合には、交換または修理されたHDD装置の固有情報が一致した第1の不揮発性メモリが存在する論理ドライブに2カ所以上の障害HDD装置が存在しなければ復旧を開始し、2カ所以上の障害HDD装置が存在する場合には、復旧は行わないようにし、複数のHDD装置のいずれかに障害が発生したとき、HDD装置の復旧が開始されたとき、及びHDD装置の復旧が終了したときに、第2の不揮発性メモリ内の管理情報を書き換え、HDD装置から読み出した固有情報を第1の不揮発性メモリへ書き込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合には当該HDD装置を障害HDD装置として処理し、管理情報を第2の不揮発性メモリに書き込む際に、書き込まれた内容と書き込んだ内容とを比較し、そ

れらが一致していない場合にはディスクアレイ装置障害として処理する。

【0028】本発明の第4のディスクアレイ装置制御方法は、上述の第3のディスクアレイ装置制御装置においてHDD装置ごとに独立に設けられた第1の不揮発性メモリに固有情報を格納する代わりに、各HDD装置ごとにその固有情報を、各HDD装置にそれぞれ対応する領域を有する単一の第1の不揮発性メモリにそれぞれ格納するようにし、また、HDD装置から読み出した固有情報を第1の不揮発性メモリへ書き込む際に書き込まれた内容

と書き込んだ内容が一致していない場合にはディスクアレイ装置障害として処理するようにしたものである。

【0029】本発明の第1のディスクアレイ装置は、データを冗長構成にして格納するための複数のHDD装置と、各HDD装置ごとに設けられ、それぞれが対応するHDD装置の識別に用いる固有情報を格納する、複数の第1の不揮発性メモリと、複数のHDD装置の状態を一括して管理するために用いる管理情報を格納するための第2の不揮発性メモリと、ディスクアレイ装置の運用中に、故障したHDD装置が交換または修理されてそのHDD装置の電源が再投入されたとき、交換または修理されたHDD装置から読み出された固有情報と、交換または修理されたHDD装置に対応する第1の不揮発性メモリに格納されている固有情報とを比較し、それらが一致する場合には復旧を開始し、それらが一致しない場合には交換または修理されたHDD装置から固有情報を読み出すとともに、読み出した固有情報を交換または修理されたHDD装置に対応する第1の不揮発性メモリに書き込み、復旧を開始するHDD装置自動復旧開始手段と、HDD装置から固有情報を読み出し、当該HDD装置に対応する第1の不揮発性メモリに書き込む固有情報書込手段と、複数のHDD装置のいずれかに障害が発生したとき、HDD装置の復旧が開始されたとき、HDD装置の復旧が終了したときに、管理情報の内容を書き換える管理情報書換手段と、HDD装置から読み出した固有情報を第1の不揮発性メモリへ書き込む際に、書き込まれた内容

と書き込んだ内容とを比較し、それらが一致していない場合には当該HDD装置を障害HDD装置として処理する固有情報書込障害検出手段と、管理情報を第2の不揮発性メモリに書き込む際に、書き込まれた内容と書き込んだ内容を比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する管理情報書込障害検出手段と、を具備する。

【0030】本発明の第2のディスクアレイ装置は、上述の第1のディスクアレイ装置において、第1の不揮発性メモリとして、複数のHDD装置に対して共通に設けられ、それぞれのHDD装置ごとに対応する領域を有する単一の第1の不揮発性メモリを用い、各HDD装置から読み出された固有情報はこの第1の不揮発性メモリにおける対応する領域に格納されるようにし、さらに、固

有情報書込障害検出手段の代わりに、HDD装置から読み出した固有情報を第1の不揮発性メモリへ書き込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する不揮発性メモリ情報書込障害検出手段を設けたものである。

【0031】本発明の第3のディスクアレイ装置は、上述の第1のディスクアレイ装置において、それぞれの冗長を構成するHDD装置のまとまりとしての論理ドライブが設定されるようにしたものであり、HDD装置自動復旧開始手段の代わりに、ディスクアレイ装置の運用中に、故障したHDD装置が交換または修理されてそのHDD装置の電源が再投入されたとき、交換または修理されたHDD装置から読み出される固有情報と全ての第1の不揮発性メモリに格納されている固有情報とを比較し、一致するものがなければ復旧を開始し、一致するものがあれば各論理ドライブに2カ以上の障害HDD装置がないことを確認し、各論理ドライブに障害HDD装置が2カ以上ない場合には復旧を開始し、論理ドライブのいずれかに2カ以上の障害HDD装置がある場合には、交換または修理されたHDD装置の固有情報が一致した第1の不揮発性メモリが存在する論理ドライブに2カ以上の障害HDD装置が存在しなければ復旧を開始し、2カ以上の障害HDD装置が存在する場合には、復旧は行わないようにする交換HDD装置復旧開始判定手段を設けたものである。

【0032】本発明の第4のディスクアレイ装置は、上述の第3のディスクアレイ装置において、第1の不揮発性メモリとして、複数のHDD装置に対して共通に設けられ、それぞれのHDD装置ごとに対応する領域を有する単一の第1の不揮発性メモリを用い、各HDD装置から読み出された固有情報はこの第1の不揮発性メモリにおける対応する領域に格納されるようにし、さらに、固有情報書込障害検出手段の代わりに、HDD装置から読み出した固有情報を第1の不揮発性メモリへ書き込む際に、書き込まれた内容と書き込んだ内容とを比較し、それらが一致していない場合にはディスクアレイ装置障害として処理する不揮発性メモリ情報書込障害検出手段を設けたものである。

【0033】すなわち上述の第1のディスクアレイ装置制御方法及び第1のディスクアレイ装置では、ディスクアレイ装置運用中に障害HDD装置の電源が再投入された際に、電源が再投入されたHDD装置の固有情報と、そのHDD装置に対応する第1の不揮発性メモリ内の固有情報とを比較し、それらが一致する場合は、復旧を開始し、それらが一致しない場合には交換または修理されたHDD装置から固有情報を読み出すとともに、その固有情報をHDD装置に対応する第1の不揮発性メモリに書き込み復旧を開始する。固有情報書込手段は、HDD装置の識別に用いる固有情報、例えばベンダー名とシリ

アル番号をHDD装置から読み出し、そのHDD装置に対応する第1の不揮発性メモリ内に書き込む。管理情報書換手段は、内蔵するHDD装置のいずれかに障害が発生した際やHDD装置の復旧が開始された際、またはHDD装置の復旧が終了した際に、障害HDD装置を縮退（切り離し）させるため、障害HDD装置の位置や復旧中のHDD装置の位置を特定できるような管理情報、例えばそれぞれのHDD装置を2ビットに対応させ“00”だと正常動作で“01”だと障害発生、“11”だと復旧中とし、第2の不揮発性メモリに書き込む。固有情報書込障害検出手段は、それぞれHDD装置に対応する第1の不揮発性メモリに固有情報を書き込みを行った際に、書き込まれた情報が間違っていないか、書き込みを行った第1の不揮発性メモリから固有情報を読み出して書き込んだ情報と比較し、それらが一致しない場合には、固有情報を格納するための不揮発性メモリが障害を起こしているため、固有情報の確認ができないので障害HDD装置として切り離し、障害を起こした不揮発性メモリ使用するHDD装置を使用しないようにする。管理情報書込障害検出手段は、ディスクアレイ装置の管理情報を第2の不揮発性メモリへ書き込みに行った際に、書き込まれた情報が間違っていないか、第2の不揮発性メモリから管理情報を読み出して書き込んだ情報と比較し、それらが一致しない場合には、第2の不揮発性メモリが障害を起こしているためディスクアレイ装置の再立ち上げのとき、正確な状態の振り分けができなくなるので、ディスクアレイ装置を障害ディスクアレイ装置として報告する。

【0034】すなわち第2のディスクアレイ装置制御方法及び第2のディスクアレイ装置では、上述の第1のディスクアレイ装置制御方法及び第1のディスクアレイ装置において、固有情報を単一の第1の不揮発性メモリにおける該当するHDD装置に対応する領域内に書込む。また、不揮発性メモリ情報書込障害検出手段は、第1の不揮発性メモリのそれぞれのHDD装置に対応した領域へ固有情報を書き込みを行った際に、書き込まれた情報が間違っていないか、書き込みを行った不揮発性メモリの領域から固有情報を読み出して書き込んだ情報と比較し、それらが一致しない場合には、固有情報を格納するための第1の不揮発性メモリが障害を起こしているため固有情報の確認ができなくなり、さらに他のHDD装置に対応した領域に関しても、同様の障害が起きることが考えられるため、ディスクアレイ装置をディスクアレイ装置障害として報告する。

【0035】すなわち第3のディスクアレイ装置制御方法及び第3のディスクアレイ装置では、ディスクアレイ装置運用中に障害HDD装置の電源が再投入された際に、電源が再投入されたHDD装置の固有情報と、全てのHDD装置に対応する不揮発性メモリ内の固有情報とを比較し、それらが一致しなければ、電源の再投入が行

われたHDD装置は交換されて新しくなっているので、固有情報を電源が再投入されたHDD装置に対応する第1の不揮発性メモリに書き込み、復旧を開始する。それらが一致する場合には、各論理ドライブに2カ以上の障害HDD装置がないことを確認し、各論理ドライブに障害HDD装置が2カ以上ない場合、電源の再投入が行われたHDD装置は、修理されてもとの位置に搭載されたか、あるいは他の論理ドライブからHDD装置を抜き取り別の論理ドライブに搭載したと判断でき、これらの場合、冗長構成によって全ての論理ドライブが復旧可能な状態であるので、固有情報が一致した第1の不揮発性メモリと電源の再投入が行われたHDD装置が一致する場合には復旧を開始し、固有情報が一致した不揮発性メモリと電源の再投入が行われたHDD装置が一致しない場合には、固有情報を電源が再投入されたHDD装置に対応する第1の不揮発性メモリに書き込み、復旧を開始する。障害HDD装置が2カ以上存在する論理ドライブがある場合で、電源の再投入が行われたHDD装置の固有情報が一致した第1の不揮発性メモリが存在する論理ドライブに2カ以上の障害HDD装置が存在しない場合、冗長構成によって復旧可能な状態であるので、固有情報が一致した第1の不揮発性メモリと電源の再投入が行われたHDD装置が一致する場合には復旧を開始し、固有情報が一致した不揮発性メモリと電源の再投入が行われたHDD装置が一致しない場合には、固有情報を電源が再投入されたHDD装置に対応する第1の不揮発性メモリに書き込み、復旧を開始する。さらに、障害HDD装置が2カ以上存在する論理ドライブがある場合で、電源の再投入が行われたHDD装置の固有情報が一致した不揮発性メモリが存在する論理ドライブに2カ以上の障害HDD装置が存在する場合、冗長構成を形成するデータを破壊する恐れがあるため、上位装置に報告を行う。固有情報書込手段、管理情報書換手段、固有情報書込障害検出手段及び管理情報書込障害検出手段の働きについては、上述の第1のディスクアレイ装置制御方法及び第1のディスクアレイ装置の場合と同様である。

【0036】すなわち第4のディスクアレイ装置制御方法及び第4のディスクアレイ装置では、上述の第3のディスクアレイ装置制御方法及び第3のディスクアレイ装置において、固有情報を単一の第1の不揮発性メモリにおける該当するHDD装置に対応する領域内に書込む。また、不揮発性メモリ情報書込障害検出手段は、第1の不揮発性メモリのそれぞれのHDD装置に対応した領域へ固有情報を書き込みを行った際に、書き込まれた情報が間違っていないか、書き込みを行った不揮発性メモリの領域から固有情報を読み出して書き込んだ情報と比較し、それらが一致しない場合には、固有情報を格納するための第1の不揮発性メモリが障害を起こしているため固有情報の確認ができなくなり、さらに他のHDD装置

に対応した領域に関しても、同様の障害が起きることが考えられるため、ディスクアレイ装置をディスクアレイ装置障害として報告する。

【0037】本発明のディスクアレイ装置制御方法及びディスクアレイ装置によれば、ディスクアレイ装置運用中に、障害HDD装置の交換または修理を行い、HDD装置の電源を再投入した際に、電源の再投入が行われたHDD装置の固有情報と不揮発性メモリの固有情報を比較することにより、HDD装置交換時のミスによるデータの損失を妨げる。

【0038】

【発明の実施の形態】次に、本発明の好ましい実施の形態について、図面を参照して説明する。図1は、本発明の実施の一形態のディスクアレイ装置の構成を示すブロック図である。

【0039】図1に示すディスクアレイ装置11は、上位装置21との接続に用いられるインターフェース12と、ディスクアレイコントローラ13と、 n 台のHDD装置14₁～14 _{n} と、各HDD装置14₁～14 _{n} に対応した不揮発性メモリ15₁～15 _{n} と、管理情報を格納する不揮発性メモリ19と、を備えている。上位装置21からのこのディスクアレイ装置11へのアクセス要求（読み出しまたは書き込み要求）は、インターフェース12を介してディスクアレイコントローラ13に入力する。

【0040】ディスクアレイコントローラ13は、上位装置21からの要求内容に応じて各HDD装置14₁～14 _{n} を制御しデータの読み出しまたは書き込みを行う機構と、不揮発性メモリ19内の情報を基にHDD装置14₁～14 _{n} の状態の判定を行う機構と、不揮発性メモリ15₁～15 _{n} 内の情報とHDD装置14₁～14 _{n} の固有情報20₁～20 _{n} を基にHDD装置の状態を判定する機構と、を有する。なお、ディスクアレイコントローラ13のこれらの機構は、制御用プロセッサと、その制御用プロセッサを規定するプログラムが格納されたプログラムメモリと、動作時に使用するパラメータを格納するメモリと、各HDD装置14₁～14 _{n} に読み出しまたは書き込みを行うためのインターフェースと、HDD装置14₁～14 _{n} にそれぞれ対応した不揮発性メモリ15₁～15 _{n} へのインターフェースと、管理情報を格納する不揮発性メモリ19へのインターフェースとによって、構成されている。固有情報20₁～20 _{n} は、それぞれ、HDD装置14₁～14 _{n} 内に格納されるものである。

【0041】このディスクアレイ装置11は、1あるいは複数の論理ドライブを設定できるものである。論理ドライブとは、それぞれの冗長を構成するHDD装置のまとまりとして設定されるものである。

【0042】まず、図2のフローチャートを用い、論理ドライブが1個である場合のこのディスクアレイの動作について、ディスクアレイ装置の運用中に1つのHDD

装置の電源が再投入された場合の処理により、説明する。

【0043】このディスクアレイ装置では、各HDD装置や不揮発性メモリなどを管理するために、カウンタ i やデータ配列 $A(i)$ 、 $B(i)$ 、 $C(i)$ 、 $D(i)$ などのパラメータを使用する。カウンタ i はHDD装置を識別するためのパラメータであり、HDD装置の固有情報を格納する固有情報配列を $A(i)$ 、不揮発性メモリに固有情報を格納する固有情報配列を $B(i)$ 、障害HDD装置の管理情報を格納する管理情報配列を $C(i)$ 、復旧を行うHDD装置の管理情報を格納する管理情報配列を $D(i)$ とする。これらのパラメータの一時的な格納には、ディスクアレイコントローラ13内のメモリが用いられる。またディスクアレイコントローラ13では、上述の制御用プロセッサにより、HDD装置14₁～14 _{n} の自動復旧を開始するためのHDD装置自動復旧開始手段と、HDD装置14₁～14 _{n} から固有情報を読み出して不揮発性メモリ15₁～15 _{n} に書き込む固有情報書き込み手段と、不揮発性メモリ19内に格納されている管理情報を書き換える管理情報書き換え手段と、不揮発性メモリ15₁～15 _{n} への固有情報書き込み時の障害を検出する固有情報書き込み障害検出手段と、不揮発性メモリ19への管理情報書き込み時の障害を検出する管理情報書き込み障害検出手段とが構成されている。

【0044】ディスクアレイ装置の運用中にHDD装置の電源の再投入があった場合、ディスクアレイコントローラ13は、まず、電源の再投入があったHDD装置の識別番号をパラメータ N_{NEW} にセットし、また、カウンタ i に“1”をセットする（ステップ101）。そして、 N_{NEW} 番のHDD装置（電源の再投入があったHDD装置）から固有情報 $A(N_{NEW})$ を読み出し（ステップ102）、この読み出しが正常に行えたかどうかを判別する（ステップ103）。 N_{NEW} 番のHDD装置から固有情報 $A(N_{NEW})$ の読み出しを正常に行えなかった場合は、ステップ112へ移る。固有情報 $A(N_{NEW})$ を正常に読み出した場合には、 N_{NEW} 番の不揮発性メモリ（不揮発性メモリ15₁～15 _{n} のうち該当するもの）から固有情報 $B(N_{NEW})$ を読み出し（ステップ104）、不揮発性メモリ15からの読み出しが正常に行えたかどうかを判別する（ステップ105）。

【0045】ステップ105において、 N_{NEW} 番の不揮発性メモリから固有情報 $B(N_{NEW})$ を正常に読み出せなかった場合には、ステップ112へ移る。一方、 N_{NEW} 番の不揮発性メモリから固有情報 $B(N_{NEW})$ を正常に読み出した場合には、 N_{NEW} 番のHDD装置の固有情報 $A(N_{NEW})$ と N_{NEW} 番の不揮発性メモリの固有情報 $B(N_{NEW})$ を比較し（ステップ106）、両者が一致する場合にはステップ109へ進む。ステップ106において $A(N_{NEW})$ と $B(N_{NEW})$ が一致しない場合には、 N_{NEW} 番の不揮発性メモリの固有情報 $B(N_{NEW})$ に N_{NEW} 番のH

DD装置の固有情報A(NNEW)を入れ、NNEW番の不揮発性メモリの固有情報B(NNEW)を書き込み(ステップ107)、この書き込みが正常に行えたかどうかを判別する(ステップ108)。NNEW番の不揮発性メモリへの固有情報B(NNEW)の書き込みが正常に行えた場合には、ステップ109へ進み、この書き込みが正常に行えなかった場合には、ステップ112へ移る。ここで書き込みが正常に行えなかったとき、NNEW番のHDD装置の状態は電源再投入前、すなわち、障害HDD装置の状態である。

【0046】ステップ109では、復旧を行うNNEW番のHDD装置の復旧状態を示す管理情報D(NNEW)に“1”を入れ、管理情報を格納する不揮発性メモリ19へ書き込む。そして、この書き込みが正常に行えたかどうかを判別する(ステップ110)。管理情報を格納する不揮発性メモリ19へ復旧状態を示す管理情報D(NNEW)を正常に書き込めなかった場合には、ステップ112へ移る。一方、ステップ110において正常に書き込みが行えた場合には、ステップ111へ移り復旧を開始する。

【0047】ステップ112では、エラーメッセージを報告し、処理を終了する。このとき、NNEW番のHDD装置からの固有情報A(NNEW)の読み出しに失敗した場合や、NNEW番の不揮発性メモリへの読み出し、書き込みに失敗した場合は、NNEW番のHDD装置の状態が電源再投入前、すなわち、障害HDD装置の状態であり、管理情報D(NNEW)の不揮発性メモリ19への書き込みが失敗した場合には、ディスクアレイ装置障害とする。エラーメッセージの報告方法は、上位装置21へのステータス報告やLED表示等があるが、障害HDD装置の位置やディスクアレイ装置障害が分かる方法であればよい。

【0048】このディスクアレイ装置では、不揮発性メモリ19へ書き込む障害HDD装置の状態を示す管理情報C(i)、復旧を示すHDD装置の状態を示す管理情報をD(i)として、障害HDD装置を示すビットと復旧中のHDD装置であることを示すビットとをそれぞれHDD装置に対応させて用いている。HDD装置が全て正常に動作しているとき、障害HDD装置を示す管理情報C(i)のビットおよび復旧を示す管理情報D(i)のビットは全て“0”であり、障害HDD装置がある場合はその障害HDD装置に対応する管理情報C(i)のビットが“1”、復旧を示す管理情報D(i)のビットは全て“0”である。また、復旧を行っているHDD装置は、復旧中のHDD装置を示す管理情報D(i)のビットが“1”であるが、このとき復旧中を示すビットが“1”になっているものは1つだけである。

【0049】また、このディスクアレイ装置では、HDD装置の固有情報A(i)及び不揮発性メモリ15の固有情報B(i)として、ベンダー名とシリアル番号を含む3

6バイトのInquiryデータを用いている。固有情報はこれに限られるものではなく、例えば、乱数を用いることや、スロット番号を用いることや、日付けや時間を示す情報を用いることができるが、HDD装置ごとに、対応する固有情報が確実に異なるものを使用することが望ましい。

【0050】次に、図3及び図4のフローチャートを用い、論理ドライブが2つ以上ある場合におけるこのディスクアレイの動作について、ディスクアレイ装置の運用中に1つのHDD装置の電源が再投入された場合の処理により、説明する。

【0051】図3及び図4は、全体として1枚のフローチャートを表すものであり、丸付きの符号P、Qによって相互の接続関係が示されていて、ディスクアレイコントローラ13の動作の流れを示している。

【0052】ここでは、各HDD装置や不揮発性メモリなどを管理するために、カウンタi、j、NJ、NEQUALやデータ配列A(i)、B(i)、C(i)、D(i)、E(i)などのパラメータを使用する。カウンタiはHDD装置を識別するためのパラメータであり、カウンタjは装置内の論理ドライブを識別するパラメータであり、カウンタjは装置構成条件により、HDD装置を識別するために用いるカウンタiと関係づけられている。HDD装置の固有情報を格納する固有情報配列をA(i)、不揮発性メモリに固有情報を格納する固有情報配列をB(i)、障害HDD装置の管理情報を格納する管理情報配列をC(i)、復旧を行うHDD装置の管理情報を格納する管理情報配列をD(i)、電源の再投入が行われたHDD装置の固有情報と不揮発性メモリに格納されている固有情報が一致したHDD装置の番号を識別するために用いる配列をE(i)とする。さらに、j番の論理ドライブに存在する障害HDD装置の数を調べるための用いるカウンタをNJ、テンポラリカウンタをNEQUALとする。これらのパラメータの一時的格納には、ディスクアレイコントローラ13内のメモリが用いられる。またディスクアレイコントローラ13では、上述の制御用プロセッサにより、HDD装置141～14nの自動復旧を開始するための交換HDD装置復旧開始判定手段と、HDD装置141～14nから固有情報を読み出して不揮発性メモリ151～15nに書き込む固有情報書込手段と、不揮発性メモリ19内に格納されている管理情報を書き換える管理情報書換手段と、不揮発性メモリ151～15nへの固有情報書き込み時の障害を検出する固有情報書込障害検出手段と、不揮発性メモリ19への管理情報書き込み時の障害を検出する管理情報書込障害検出手段とが構成されている。

【0053】ディスクアレイ装置の運用中にHDD装置の電源の再投入があった場合、ディスクアレイコントローラ13は、まず、電源の再投入があったHDD装置の識別番号をNNEWにセットし、HDD装置識別カウンタ

iに“1”、電源の再投入が行われたHDD装置の固有情報と不揮発性メモリに格納されている固有情報が一致したHDD装置の番号を識別するために用いる配列E(i)に“0”、j番の論理ドライブに存在する障害HDD装置の数を調べるために用いるカウンタNjに

“0”、テンポラリカウンタNEQUALの“0”をセットする(ステップ201)。そして、N_{NEW}番のHDD装置からその固有情報A(N_{NEW})を読み出し(ステップ202)、この読み出しが正常に行えたかどうかを判別する(ステップ203)。ここで読み出しが正常に行えなかった場合には、HDD装置障害としてエラーメッセージを上位装置21に報告し(ステップ317)、処理を終了する。N_{NEW}番のHDD装置から固有情報A(N_{NEW})の読み出しが正常に行えた場合には、i番の不揮発性メモリ15から固有情報B(i)を読み出し(ステップ204)、この不揮発性メモリ15からの読み出しが正常に行えたかどうかを判別する(ステップ205)。

【0054】ステップ205においてi番の不揮発性メモリ15から固有情報B(i)の読み出しが正常に行えなかった場合には、ステップ210へ移る。一方、ステップ205において読み出しが正常に行えた場合には、電源の再投入が行われたHDD装置の固有情報A(N_{NEW})とi番の不揮発性メモリに格納されている固有情報B(i)を比較し(ステップ206)、それらが一致する場合には、E(i)に“1”をセット、テンポラリカウンタNEQUALに“1”を加算、j番の論理ドライブに存在する障害HDD装置の数を調べるためのカウンタNjに

“1”を加算し(ステップ216)、ステップ217へ移る。これに対し、ステップ206において、A(N_{NEW})とB(i)が一致しない場合には、不揮発性メモリ19から障害HDD装置の管理情報C(i)を読み出し(ステップ207)、この読み出しが正常に行えたかどうかを判別する。不揮発性メモリ19から障害HDD装置の管理情報C(i)の読み出しが正常に行えなかった場合には、ディスクアレイ装置障害としてエラーメッセージを上位装置12に報告し(ステップ317)、処理を終了する。ステップ208において管理情報C(i)の読み出しが正常に行えた場合、障害HDD装置を示すこの管理情報C(i)が“1”かどうかを判別し(ステップ209)、C(i)＝“1”ならばステップ215へ移る。ステップ209でC(i)≠“1”であれば、ステップ217へ進む。

【0055】ステップ210では不揮発性メモリ19から障害HDD装置の管理情報C(i)を読み出し、ステップ211においてこの読み出しが正常に行えたかどうかを判別する。ここで不揮発性メモリ19から管理情報C(i)の読み出しが正常に行えなかった場合には、ディスクアレイ装置障害としてエラーメッセージを上位装置21に報告し(ステップ317)、処理を終了する。一方、ステップ211において読み出しが正常に行えた場

合には、障害HDD装置を示す管理情報C(i)が“1”かどうかを判別し(ステップ212)、C(i)＝“1”ならばステップ215へ移る。C(i)≠“1”であれば、障害HDD装置を示す管理情報C(i)に“1”をセットして不揮発性メモリ19にこのC(i)を書き込み(ステップ213)、書き込みが正常に行えたかどうかを判別する(ステップ214)。管理情報C(i)の書き込みが正常に行えた場合には、ステップ215へ進み、正常に行えなかった場合には、ディスクアレイ装置障害としてエラーメッセージを上位装置21に報告し(ステップ317)、処理を終了する。

【0056】ステップ215では、j番の論理ドライブに存在する障害HDD装置の数を調べるためのカウンタNjに“1”を加算し、ステップ217へ進む。論理ドライブ識別カウンタjは、装置構成条件とHDD装置番号iから導き出される。

【0057】ステップ217では、上述の処理を全てのHDD装置に対して繰り返すために、HDD装置を識別するためのカウンタiに“1”を加算する。そして、加算後のカウンタiが全てのHDD装置台数以下であるかどうかを判別する(ステップ218)。iがHDD装置の台数以下ならばステップ204へ移り、カウンタiがHDD装置台数を越えたならば、ステップ301へ移る。

【0058】ステップ301では、テンポラリカウンタNEQUALが“0”かどうかを判別する。ここでNEQUAL＝“0”であればステップ312へ移る。この場合、電源が再投入されたHDD装置の固有情報A(N_{NEW})と不揮発性メモリに格納している固有情報B(i)は全て異なることになる。一方、ステップ301においてNEQUAL≠“0”ならば、ディスクアレイ装置内の論理ドライブを識別するためのカウンタjに“1”、テンポラリカウンタNEQUALに“0”をセットし(ステップ302)、HDD装置を識別するためのカウンタiに、論理ドライブjの最初のHDD装置の番号を入れ(ステップ303)、ステップ304へ移る。

【0059】ステップ304では、電源の再投入が行われたHDD装置の固有情報と不揮発性メモリに格納されている固有情報が一致したHDD装置の番号を識別するために用いる配列E(i)が“1”であるかどうかを判別する。E(i)≠“1”であれば、HDD装置を識別するカウンタiに“1”を加算して(ステップ305)、このカウンタiが論理ドライブjの最後のHDD装置を越えていないかどうかを判別する(ステップ306)。カウンタiが論理ドライブjの最後のHDD装置を越えていなければ、論理ドライブjに属するHDD装置に上記の処理を繰り返すためにステップ304に戻り、カウンタiが論理ドライブjの最後のHDD装置を越えている場合には、ステップ309へ移る。また、ステップ304においてE(i)が“1”ならば、j番の論理ドライブ

25

に存在する障害HDD装置の数Njが2以上かどうかを判別する(ステップ307)。Njが2以上であれば、テンポラリカウンタNEQUALに“1”を加算してから(ステップ308)、ステップ309に移り、Njが2未満ならばそのままステップ309へ移る。

【0060】ステップ309では、論理ドライブを識別するカウンタjに“1”を加算し、その後、ステップ310において、jが全論理ドライブ数以下かどうかを判別する。jが全論理ドライブ数以下であれば、ステップ303からの処理を繰り返し、jが全論理ドライブ数を越えているならば、テンポラリカウンタNEQUALが“0”かどうかのチェックを行う(ステップ311)。NEQUALが“0”でない場合は、NNEWの存在する論理ドライブに2カ以上の障害HDD装置があることにより復旧できない場合か、またはNNEWの存在するドライブ以外にNNEWのHDD装置の固有情報と同じ固有情報を持ち、さらに2カ以上の障害HDD装置があることによりデータの冗長構成を壊すおそれがある場合であるため、エラーメッセージを上位装置21に報告して(ステップ317)、処理を終了する。ステップ311においてテンポラリカウンタNEQUALが“0”ならば、ステップ312に移行する。

【0061】ステップ312では、NNEW番の不揮発性メモリの固有情報B(NNEW)に、NNEW番のHDD装置の固有情報A(NNEW)を入れ、NNEW番の不揮発性メモリの固有情報B(NNEW)を書き込む。そして、ステップ313において、NNEW番の不揮発性メモリの固有情報B(NNEW)の書き込みが正常に行えたかどうかを判別する。ここで書き込みが正常に行えている場合には、ステップ314に進み、NNEW番の不揮発性メモリの固有情報B(NNEW)の書き込みが正常に行えなかった場合には、エラーメッセージを上位装置21に報告して(ステップ317)、処理を終了する。

【0062】ステップ314では、復旧を行うNNEW番のHDD装置の復旧状態を示す管理情報D(NNEW)に“1”を入れ、管理情報を格納する不揮発性メモリ19へ書き込む。そして、復旧状態を示す管理情報D(NNEW)を管理情報を格納する不揮発性メモリ19へ正常に書き込めたかどうかを判別し(ステップ315)、正常に書き込めていなかった場合には、ディスクアレイ装置障害としてエラーメッセージを上位装置21に報告して(ステップ317)、処理を終了する。一方、ステップ315において正常に書き込みが行えた場合には、復旧を開始して(ステップ316)、処理を終了する。

【0063】以上説明した本実施形態のディスクアレイ装置は、固有情報を格納する不揮発性メモリを個々のHDD装置に対応して備え、それとは別に管理情報を格納する不揮発性メモリを用いた構成であるが、本発明が適用されるディスクアレイ装置の構成はこれに限定されるものではない。例えば、固有情報を格納する不揮発性メ

26

モリを1つにして、この不揮発性メモリの記憶領域を個々のHDD装置の固有情報を格納する領域に分けるようにしてもよいし、管理情報と固有情報とを全て1つの不揮発性メモリに領域を分けて格納するようにしてもよい。このよう全ての固有情報を1つの不揮発性メモリに格納するようにした場合には、上位装置に報告するエラーメッセージをディスクアレイ装置障害とする。これは、固有情報を格納する不揮発性メモリを1つにした場合、その不揮発性メモリ内の1つの領域の読み出したとしても、そのデータの信頼性が著しく低下するため、この不揮発性メモリを使用しないようにするためである。具体的には、ディスクアレイコントローラ13において、障害HDD装置として処理する固有情報書込障害検出手段の代わりに、ディスクアレイ装置障害として処理する不揮発性メモリ情報書込障害検出手段を設けるようにする。

【0064】また、上述の実施形態のディスクアレイ装置では、1台のHDD装置の復旧しか行えない冗長構成を採用しているが、障害HDD装置を識別するパラメータを復旧可能な台数に応じて増やし、障害HDD装置の台数を示すカウンタによる分岐条件による制御を実行することで、2台以上のHDD装置を復旧できる冗長構成をとることができる。また、上述のディスクアレイ装置では、障害HDD装置の電源の再投入時に、一部自動的に復旧可能なディスクに対しては復旧を行うように構成してあるが、この機構を設けずに、上位装置からの指示により、復旧作業が開始されるように構成してもよい。

【0065】

【発明の効果】以上説明したように本発明は、ディスクアレイ装置運用中にHDD装置の交換が行われた際、HDD装置の状態を自動検出するとともに不揮発性メモリの障害も検出することにより、交換されたHDD装置の誤動作が生じることがなくなるという効果がある。また、HDD装置の接続ミスも検出することにより、人為的ミスによりデータが失われることがなくなるという効果がある。

【図面の簡単な説明】

【図1】図1は、本発明の実施の一形態のディスクアレイ装置の構成を示すブロック図である。

【図2】図1のディスクアレイ装置において論理ドライブが1つのときにHDD装置の電源が再投入された時の処理を示すフローチャートである。

【図3】図1のディスクアレイ装置において論理ドライブが2つ以上のときにHDD装置の電源が再投入された時の処理を示すフローチャートである。

【図4】図1のディスクアレイ装置において論理ドライブが2つ以上のときにHDD装置の電源が再投入された

27

時の処理を示す図であって、図 3 に示す処理に引き続く処理を示すフローチャートである。

【図 5】従来のディスクアレイ装置の構成の概要を示すブロック図である。

【図 6】図 5 のディスクアレイ装置の再立上げの動作の処理を示すフローチャートである。

【図 7】図 5 のディスクアレイ装置の再立上げの動作の処理を示す図であって、図 6 に示す処理に引き続く処理を示すフローチャートである。

【図 8】図 5 のディスクアレイ装置の再立上げの動作の処理を示す図であって、図 6 に示す処理に引き続く処理を示すフローチャートである。

【図 9】図 5 のディスクアレイ装置の再立上げの動作の処理を示す図であって、図 8 に示す処理に引き続く処理を示すフローチャートである。

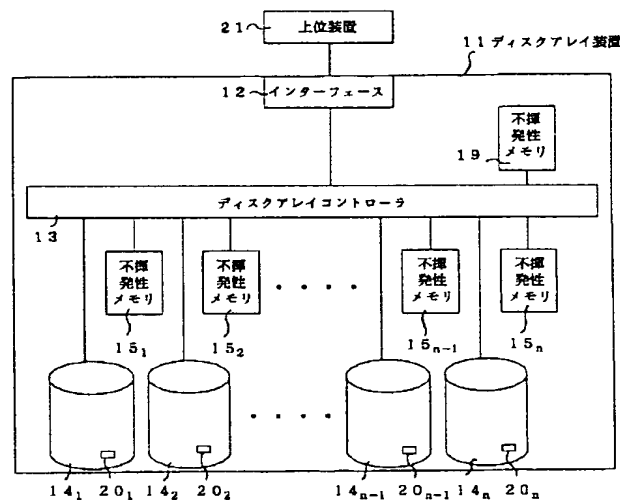
【図 10】図 5 のディスクアレイ装置の再立上げの動作の処理を示す図であって、図 9 に示す処理にから分岐する処理を示すフローチャートである。

【符号の説明】

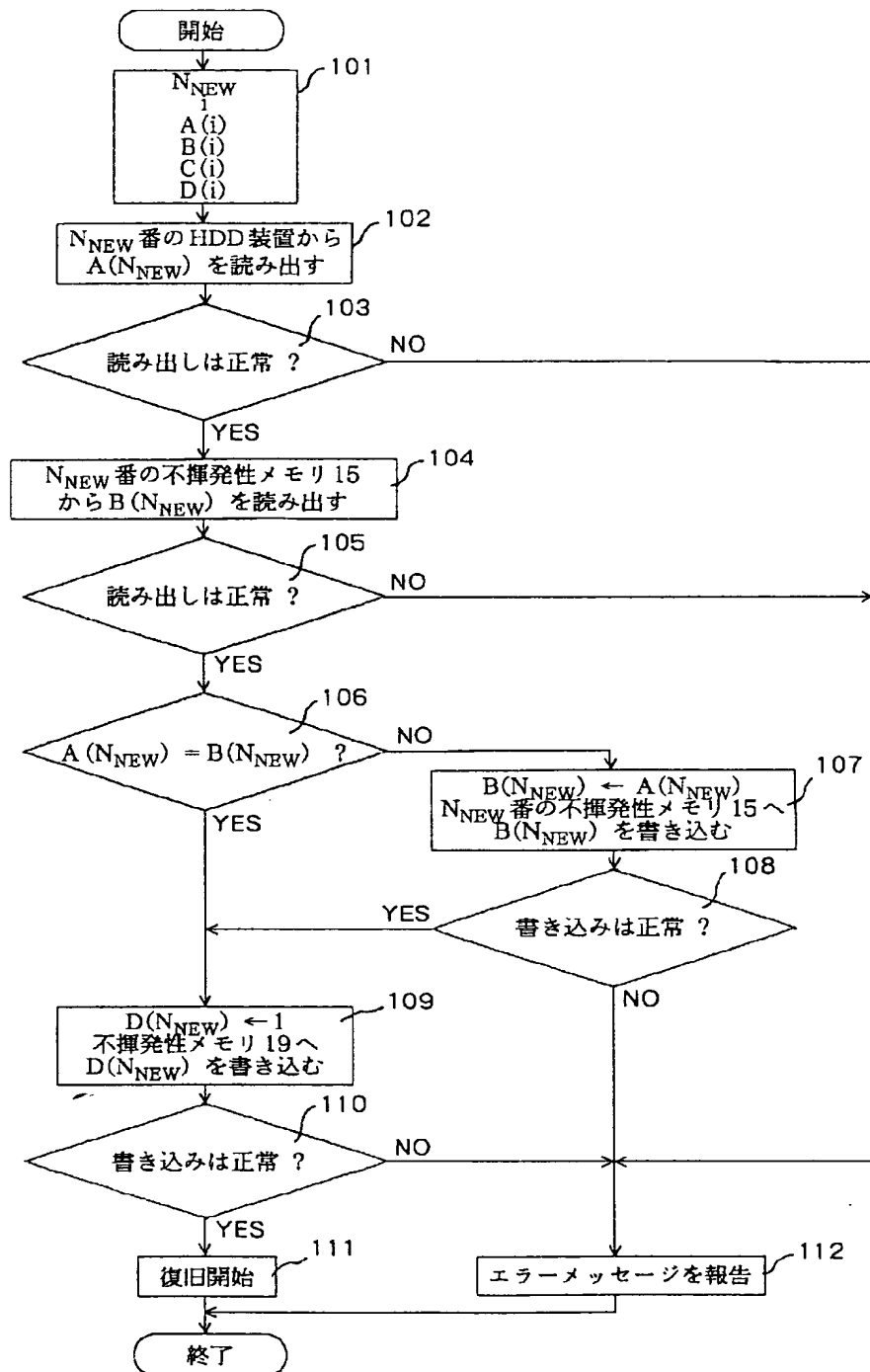
28

- 3, 13 ディスクアレイコントローラ
- 5 初期化制御手段
- 6 状態情報制御手段
- 7 直前状態際限手段
- 8 接続状態判定手段
- 9 復旧制御手段
- 11 ディスクアレイ装置
- 12 インターフェース
- 14, 14₁ ~ 14_n HDD 装置
- 15, 15₁ ~ 15_n 固有情報格納用不揮発性メモリ
- 19 管理情報格納用不揮発性メモリ (状態情報記憶部)
- 20₁ ~ 20_n 固有情報
- 21 上位装置
- 54, 54₁ ~ 54_s 磁気ディスク装置 (HDD 装置)
- 55, 55₁ ~ 55_s 固有情報記憶部 (不揮発性メモリ)

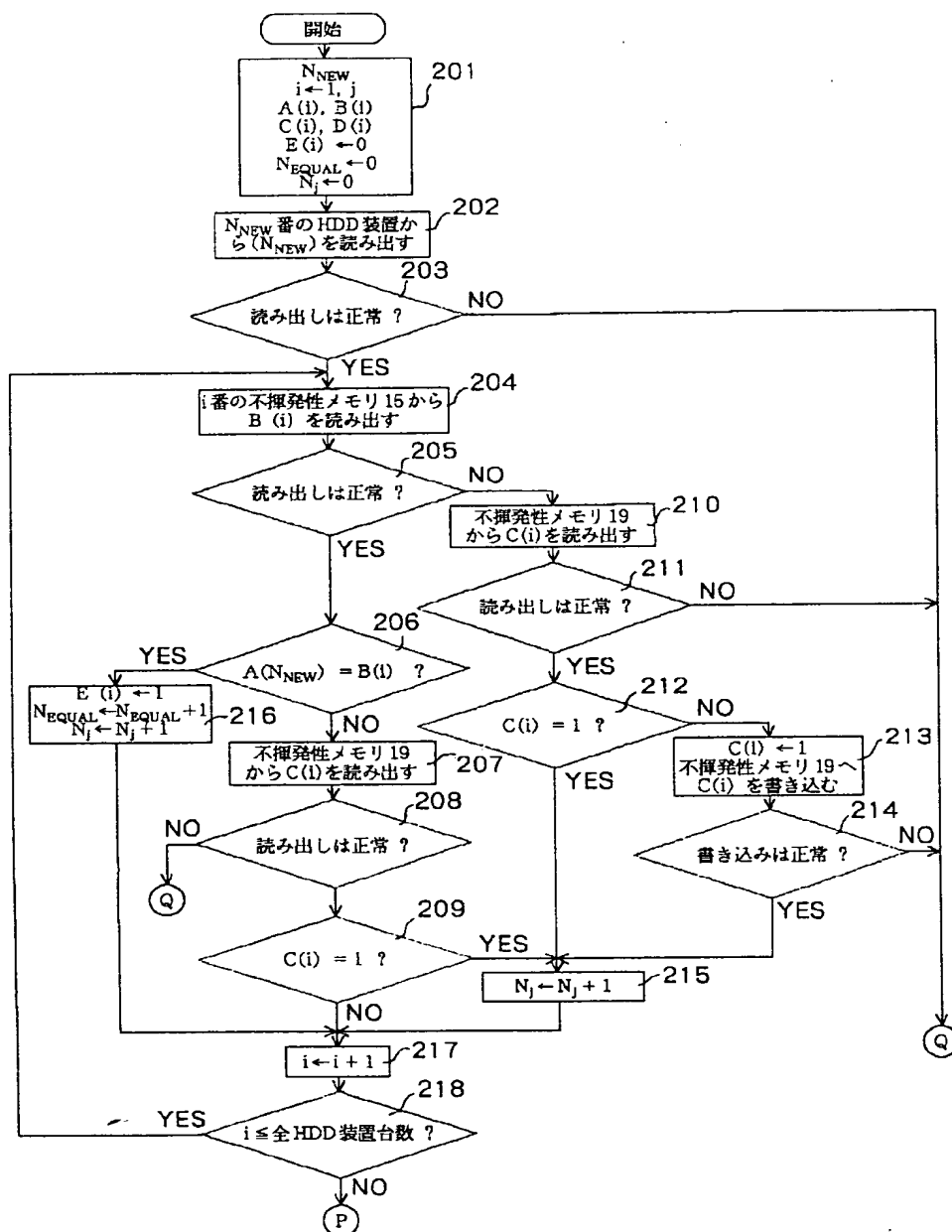
【図 1】



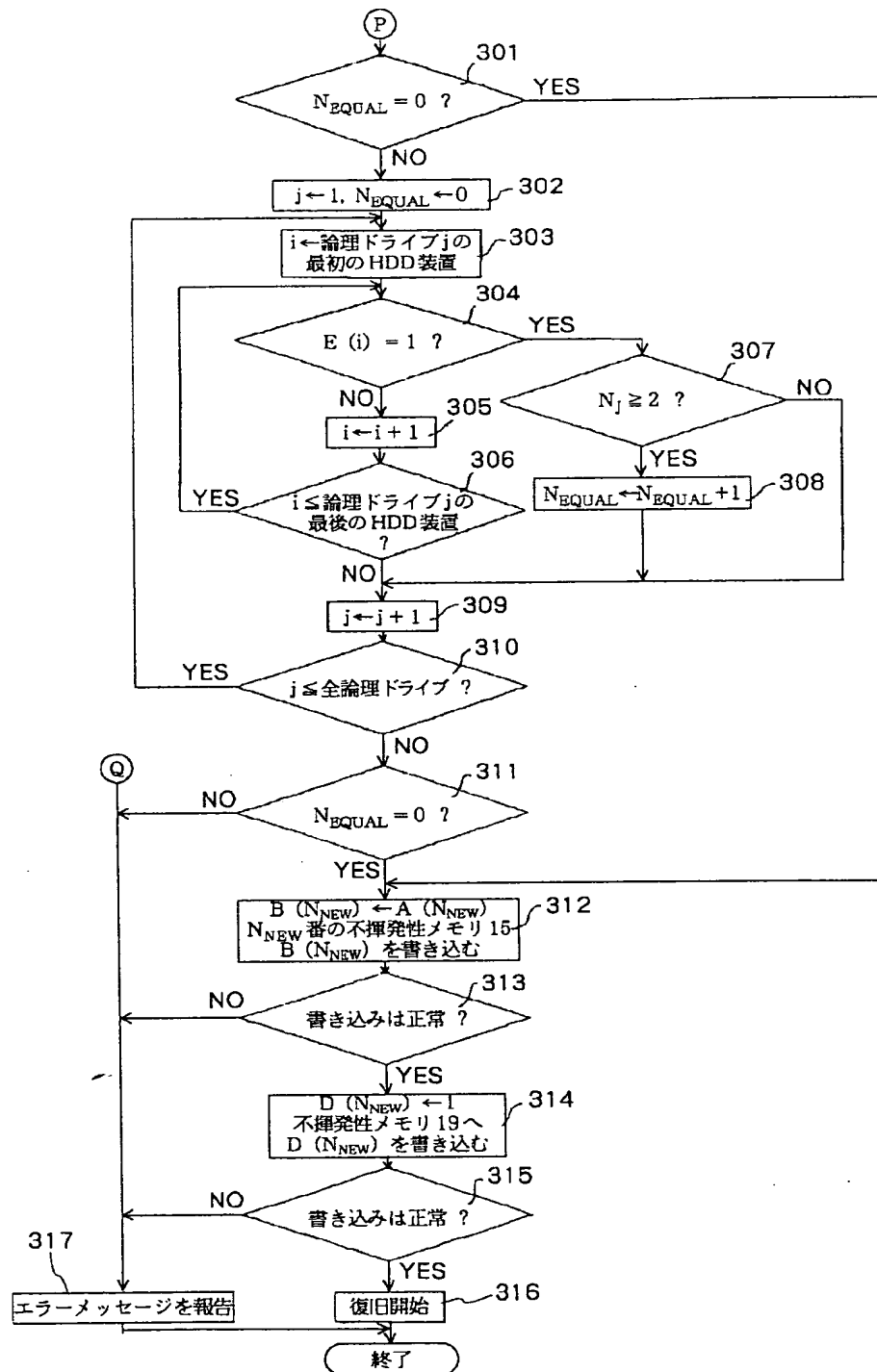
【図 2】



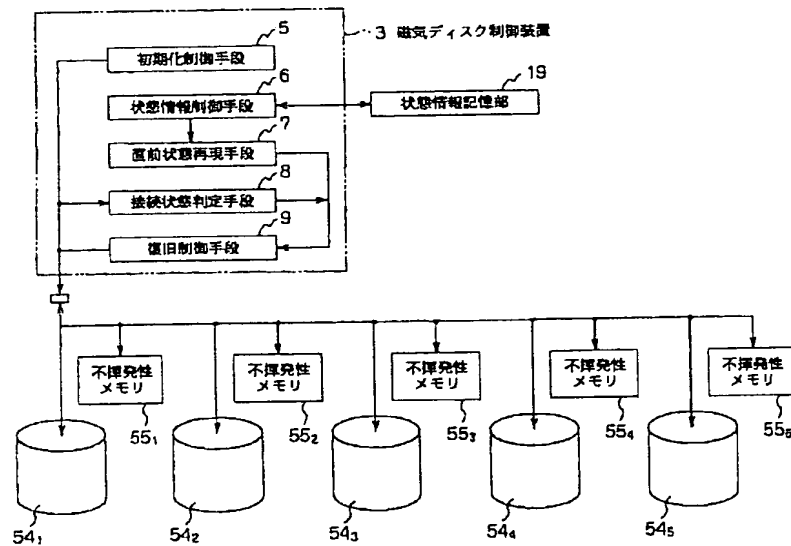
【図 3】



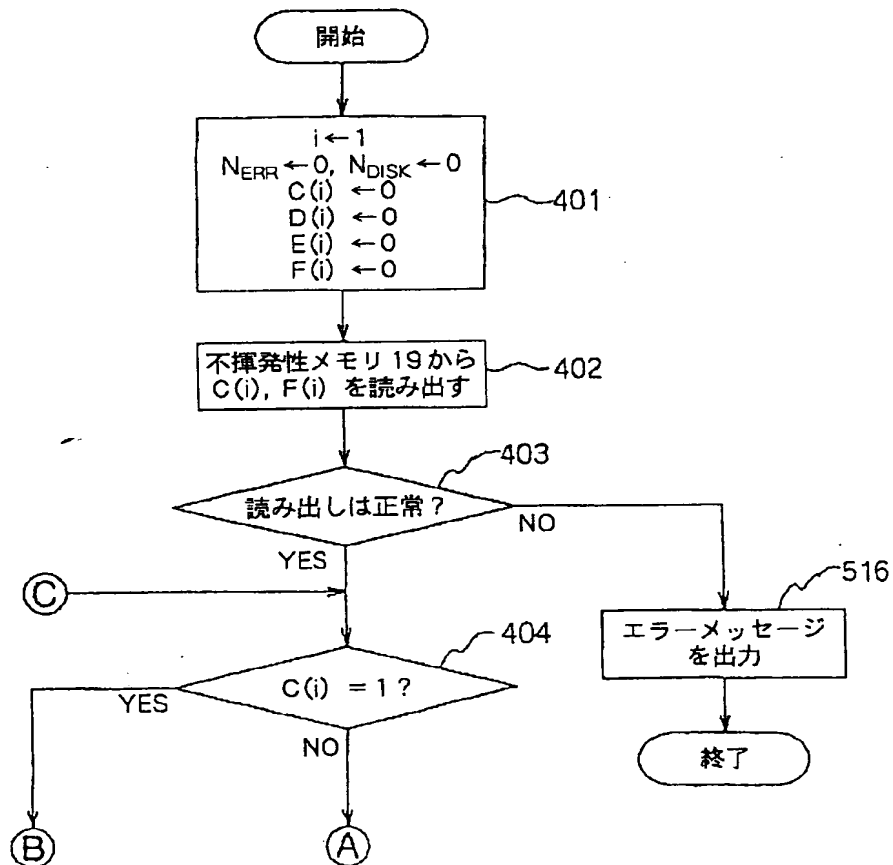
【図 4】



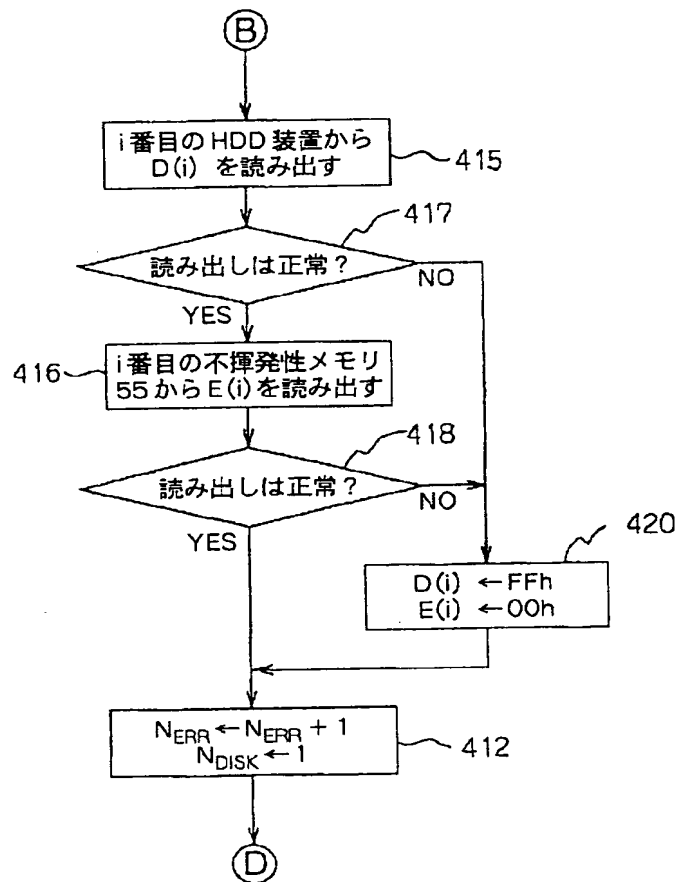
【図 5】



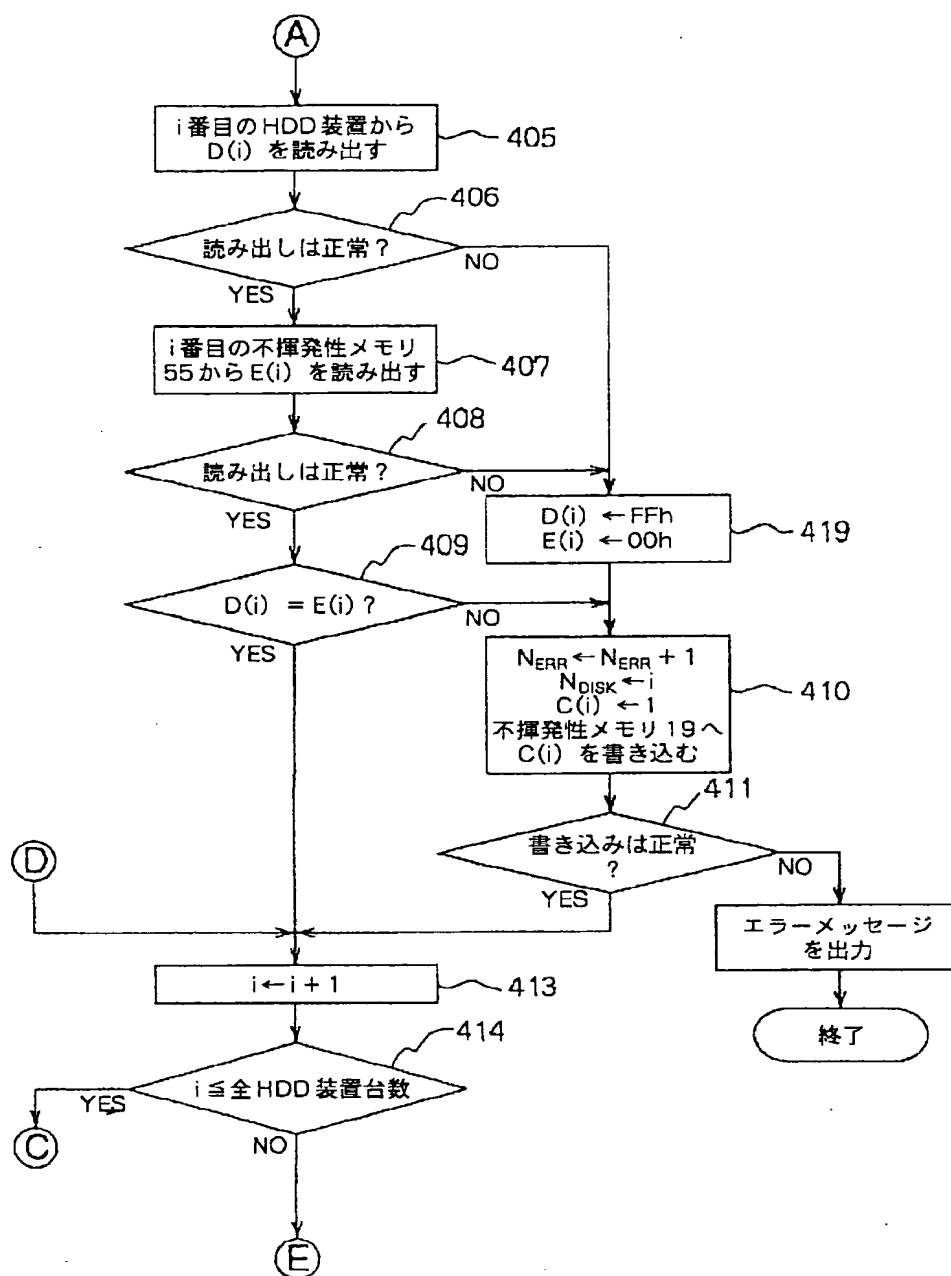
【図 6】



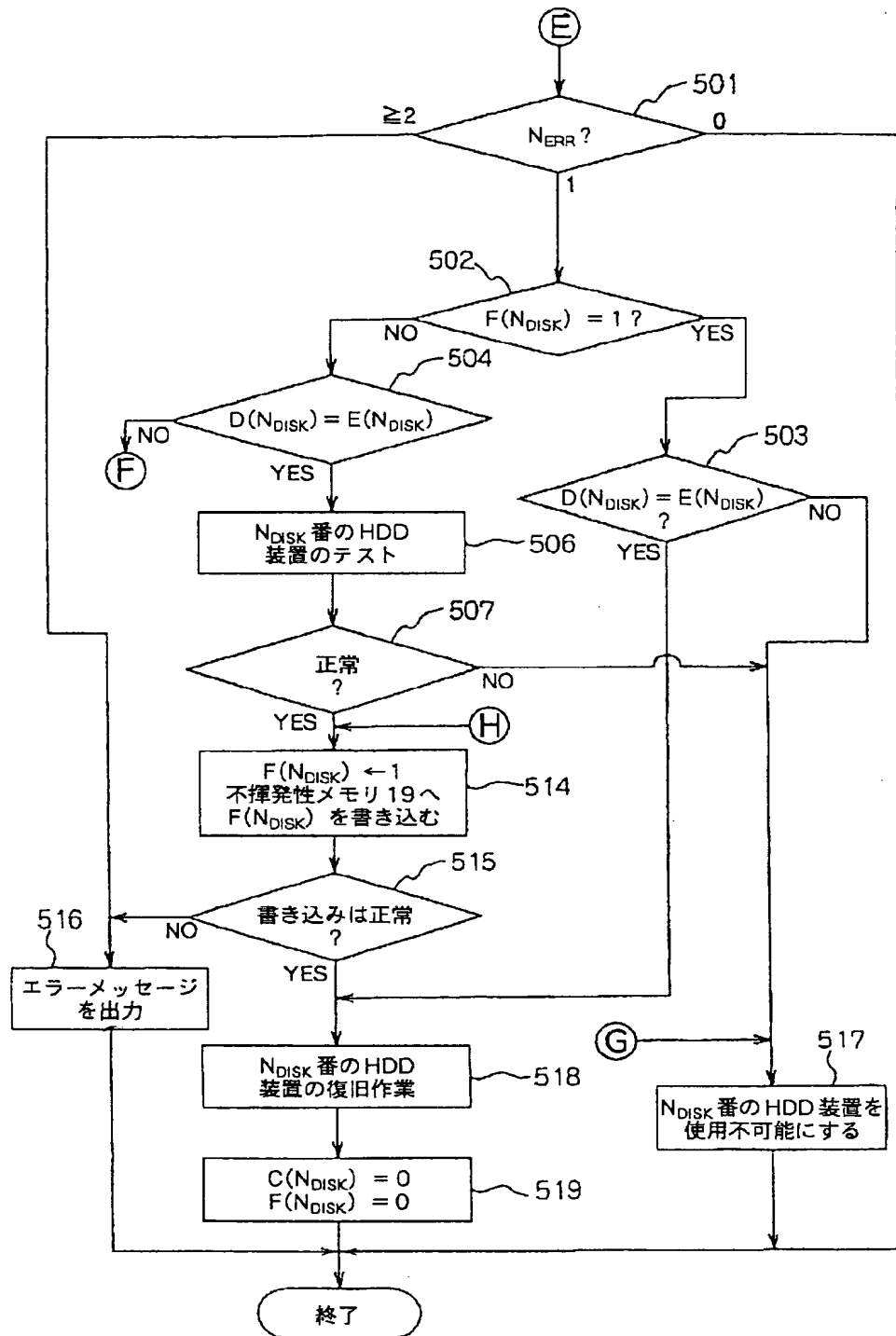
【図7】



【図8】



【図 9】



【図 10】

